Web Based Plant Leaf Identification Using Deep Learning Methodology

Monica M [1], Kaviyarasan L [1], Priyadarshini S [2]

[1]*Student, Department of Computer Science and Engineering, Adithya Institute of Technology, Coimbatore, Tamil Nadu, India.*

[2]*Assistant Professor, Department of Computer Science and Engineering, Adithya Institute of Technology, Coimbatore, Tamil Nadu, India.*

Corresponding Author: mmonica2702@gmail.com

Abstract: - In this world there are thousands of plant species available, and plants have medicinal values. Medicinal plants play a very active role in healthcare traditions. So proper identification of the medicinal plants has major benefits for not only manufacturing medicines but also for forest department peoples, life scientists, physicians, medication laboratories, government, and the public. The manual method is good for identifying plants easily, but is usually done by the skilled practitioners who have achieved expertise in this field. However, it is time consuming. There may be chances to misidentification, which leads to certain side effects and may lead to serious problems. Automatic identification and classification of different plant leaf species have become a common trend among researchers and scientists. In the field of computer vision, solving problems using deep learning algorithms has become more prevalent. In this project, we proposed **Deep med** an automated system for the medicinal plant classification, which will help people identify useful plant species quickly. The convolutional neural network is famous for its influential abilities in feature extraction and classification. The proposed model entitled **Deep med** model developed by extracting the Exception features and classified using CNN classifier shows high accuracy with less prediction time on real-time image.

Key Words: - ***Plant species, Medicinal plants, Healthcare, Automatic identification.***

I. INTRODUCTION

1.1 Overview

Plants are multicellular organisms in the kingdom Plantae that use photosynthesis to make their own food. There are over 300,000 species of plants; common examples of plants include grasses, trees, and shrubs. Plants have an important role in the world's ecosystems. They produce most of the world's oxygen, and are important in the food chain, as many organisms eat plants or eat organisms which eat plants. The study of plants is called botany.

Manuscript revised May 20, 2023; accepted May 21, 2023. Date of publication May 23, 2023.

This paper available online at www.ijprse.com

ISSN (Online): 2582-7898; SJIF: 5.59

1.2 Medicinal Plant

A medicinal plant is any plant which, in one or more of its organs, contains substances that can be used for therapeutic purposes or which are precursors for the synthesis of useful drugs. The term "medicinal plant" includes various types of plants used in herbalism ("herbology" or "herbal medicine"). It is the use of plants for medicinal purposes, and the study of such uses.

The word "herb" has been derived from the Latin word, "herba" and an old French word "herbe". Now a days, herb refers to any part of the plant like fruit, seed, stem, bark, flower, leaf, stigma or a root, as well as a non-woody plant. Earlier, the term "herb" was only applied to non-woody plants, including those that come from trees and shrubs. These medicinal plants are also used as food, flavonoid, medicine or perfume and also in certain spiritual activities.

Plants have been used for medicinal purposes long before prehistoric period. Ancient Unani manuscripts Egyptian papyrus and Chinese writings described the use of herbs.

Evidence exists that Unani Hakims, Indian Vaids and European and Mediterranean cultures were using herbs for over 4000 years as medicine. Indigenous cultures such as Rome, Egypt, Iran, Africa and America used herbs in their healing rituals, while other developed traditional medical systems such as Unani, Ayurveda and Chinese Medicine in which herbal therapies were used systematically. Treatment with medicinal plants is considered very safe as there is no or minimal side effects. These remedies are in sync with nature, which is the biggest advantage. The golden fact is that, use of herbal treatments is independent of any age groups and the sexes.

The ancient scholars only believed that herbs are only solutions to cure a number of health-related problems and diseases. They conducted thorough study about the same, experimented to arrive at accurate conclusions about the efficacy of different herbs that have medicinal value. Most of the drugs, thus formulated, are free of side effects or reactions. This is the reason why herbal treatment is growing in popularity across the globe. These herbs that have medicinal quality provide rational means for the treatment of many internal diseases, which are otherwise considered difficult to cure.

Medicinal plants such as Aloe, Tulsi, Neem, Turmeric and Ginger cure several common ailments. These are considered as home remedies in many parts of the country. It is known fact that lots of consumers are using Basil (Tulsi) for making medicines, black tea, in pooja and other activities in their day-to-day life. In several parts of the world many herbs are used to honour their kings showing it as a symbol of luck. Now, after finding the role of herbs in medicine, lots of consumers started the plantation of Tulsi and other medicinal plants in their home gardens. Medicinal plants are considered as a rich resource of ingredients which can be used in drug development either pharmacopoeia, non- pharmacopeia or synthetic drugs. A part from that, these plants play a critical role in the development of human cultures around the whole world. Moreover, some plants are considered as important source of nutrition and as a result of that they are recommended for their therapeutic values. Some of these plants include ginger, green tea, walnuts, aloe, pepper and turmeric etc. Some plants and their derivatives are considered as important source for active ingredients which are used in aspirin and toothpaste etc. Apart from the medicinal uses, herbs are also used in natural dye, pest control, food, perfume, tea and so on. In many countries different kinds of medicinal plants/ herbs are used to keep ants, flies, mice and flee away from homes and offices. Now a day's medicinal herbs are important sources for pharmaceutical manufacturing.

II. EXISTING PROBLEM

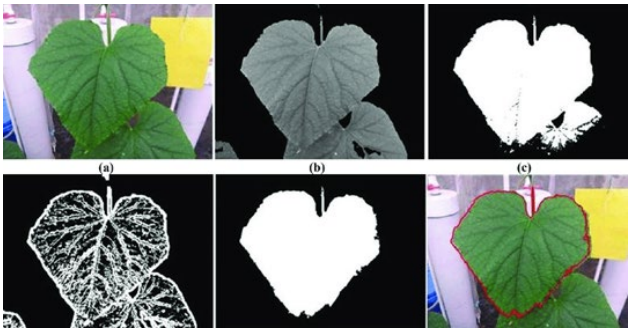
According to the State of Medicinal Plants report, there exist about 390,000 plant species known to modern science. This makes it difficult for a botanist or an expert to be able to identify and classify all these species. Even more so impossible for no experts. Not forgetting the fact that some plant species have high similarities and the differences being so little pose the need for a fine-grained classification approach. The leaves of plants contain a considerable amount of information regarding the particular plant species and a lot can be learned about a plant just from its leaves. Plants, obviously, right from the inception of creation have had a variety of use for humans and are not stopping any moment soon, ranging from their use for food, medicine, shelter, etc. The conventional method of identifying the Indian medicinal plants is time-consuming, tedious and a challenging work as it may be incorrect many a time.In addition, owing to global warming and other factors, many plants nowadays are faced with the challenge of extinction. Non-endangered as well as endangered plant species need to be preserved and conserved adequately to counter this risk. Hence, there is a need to develop an automated or computerized system to identify and classify plants in an efficient manner. Therefore, the need for adequate knowledge and management of plant species cannot be overemphasized. The traditional artificial identification methods are based onplant morphology, which though having its form of success is subjective, costlier in terms of manpower and proves inefficient in the long run. Classifying the herbs using the algorithms of computer vision show concern on categorizing the plant images into its distinct groups. The classification of plants using digital leaf images are challenging due to their similarities in inter-class and intra-class, the possibility of complex background and variations in many parameters such as illumination and color. Thus, developing tools and solutions to analyze and interpret the patterns in the leaf images with significant results are essential. A model or system to accelerate the automatic recognition of Indian medicinal herbs is yet not concentrated in recent times by many researchers. Instead of relying on skilled botanists or Ayurveda (an ancient medicinal system of India) experts, automatic identification of therapeutic herbs will spring knowledge at ease to the general public and other stakeholders of medicinal plants. Plant leaves are two-dimensional in nature; therefore, making it possible to engage in the identification of plant leaves automatically using image-processing techniques. In recent times, Deep Learning and in particular the use of Convolutional

Neural Networks (CNNs) have proven well suited for addressing computer vision problems of which plant classification can be considered to be one. Deep learning eliminates the need for domain expertise and hard-core feature extraction that only expert botanists can provide.

III. Proposed Solution

3.1 Region Proposal Network

This region proposal network takes convolution feature map that is generated by the backbone layer as input and outputs the anchors generated by sliding window convolution applied on the input feature map.



3.2 Gray Level Co-occurrence Matrix

Gray Level Co-occurrence Matrix (GLCM) based texture analysis of kidney diseases for parametric variations. The investigations were carried out using three Pyoderma variants (Boil, Carbuncle, and Impetigo Contagiosa) using GLCM. GLCM parameters (Energy, Correlation, Contrast, and Homogeneity) were extracted for each colour component of the images taken for the investigation. Contrast, correlation, energy, and homogeneity represent the coarseness, linear dependency, textural uniformity, and pixel distribution of the texture, respectively. The analysis of the GLCM parameters and their histograms showed that the said textural features are disease dependent. The approach may be used for the identification of CKD diseases with satisfactory accuracy by employing a suitable deep learning algorithm.

3.3 Convolutional Neural Network (CNN)

A CNN is a type of deep learning used to analyse visual scenes. It is characterized by having one or more hidden layers, which extract the attributes in videos or images, and a fully connected layer to produce the desired output. Whereas for the computer, the image is a 3D array (width × height × depth) of values ranging from 0 to 255. It is simply pixels of colour; if the number of channels is one, the image is grayscale, black, and white. Besides, the channels are three colours (if images are RGB). CNN Deep Network has shown outstanding performance in many competitions related to image processing due to its accurate results. CNN is a hierarchical structure that contains several layers.

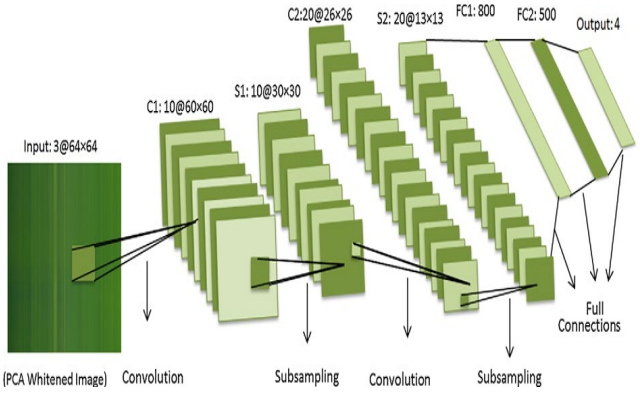


Fig.1. Architecture of CNN

The basic components of the basic convolutional neural networks are: The Convolutional Layer, the Activating function, the Pooling Layer, and the Fully-connected Layer.

Convolutional Layer: In the convolutional layer, a filter (known as a kernel) is used to determine the existence of patterns in the input images (original image), after which several filters can be employed to extract different features. The filter is a small size to have the ability to scan the whole image and apply the appropriate arithmetic between the filter and the pixels to extract the features. The filter settings are reset during the periodic training phase, and when the network has been trained for a retinacular number of epochs (epochs imply all training samples have been entered simultaneously), these filters start looking for different characteristics in the image. Simple and evident features, such as edges in various directions, are extracted using the first hidden layers. The complexity of the attributes which must be recognized and extracted rises as we go deeper into the network's hidden levels.

Pooling Layer: The purpose of the pooling is to reduce the size of the activation maps. This is not necessary but prevents you from falling into an overfitting situation. The idea behind clustering is simple, as large arrays are scaled down.
Fully-connected Layer: This layer is the last, where neurons are fully connected to all nodes of the previous layer. The final classification process takes place in it.

To design the network model, first, an image is inserted into a conv layer, and an activation function is applied to the output of the conv layer, such as ReLu. The function's output is sent to another conv layer; the process is repeated several times, sending the output to an assembly layer. The steps are repeated several times, and trainable classifiers are produced. The output is also sent to the fully connected layer, which has the probability of each class we want to train the network on. In the input layer, the range can be from 0 to 1. Each neuron is treated as a filter where the filter is computed for the data network depth; in the conv layer, the neurons are filters in image processing to detect edges, curves, etc. Each filter of the conv layer will have its image features, such as vertical edges, horizontal edges, colours, textures, and density.

All neurons add to the feature extractor array for the entire image. In addition, the pooling layer is sandwiched between successive convolutional layers to compress the amount of data and parameters and reduce overfitting. In short, if the input is an image, then the main function of the pooling layer is to compress the image by resizing the image. When the information removed when the image is compressed is just some irrelevant information, we can remove it.

IV. Final Discussions

The important points involved with the performance metrics are discussed based on the context of this project:

True Positive (TP): There is a leaf, and the algorithms detect plant name.

False Positive (FP): There is no leaf, but the algorithms detect as leaf and display plant name.

False Negative (FN): There is a leaf, but the algorithms do not detect Leaf and plant name.

True Negative (TN): There is no leaf, and nothing is being detected.

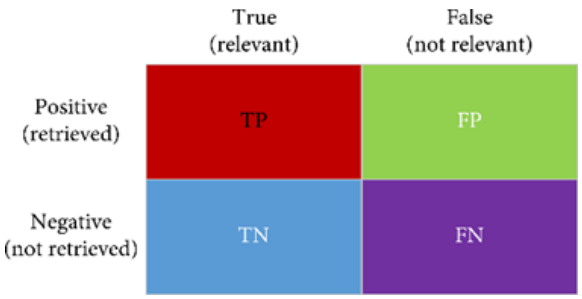
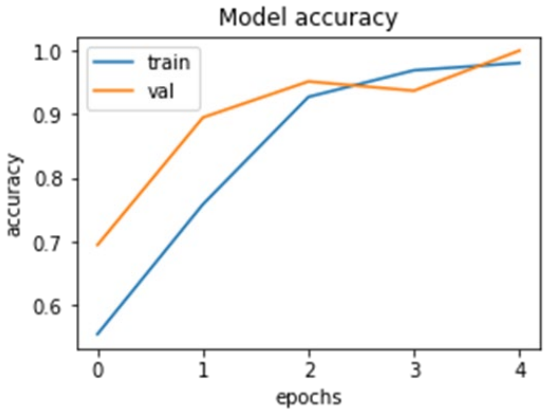


Fig.2. Parameter Definition

4.1 Accuracy

Accuracy is a measure that tells whether a model/algorithm is being trained correctly and how it performs. In the context of this thesis, accuracy tells how well it is performing in detecting plant species. Accuracy is calculated using the following formula.

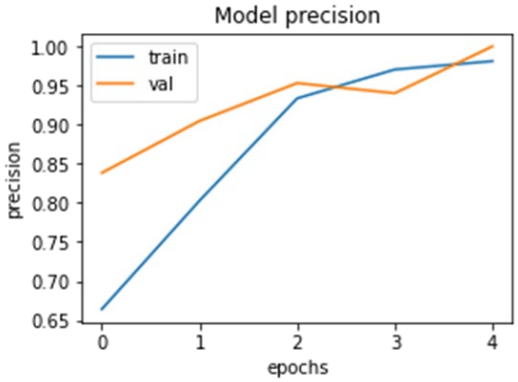
Accuracy = (T P + T N)/ (T P + T N + F P + F N)



4.2 Precision

It denotes the ratio of positively predicted cases that are actually positive. In the context of this thesis, precision measures the fraction of objects that are predicted to be plant and are actually leaf of that plant type. Precision is calculated using the following formula.

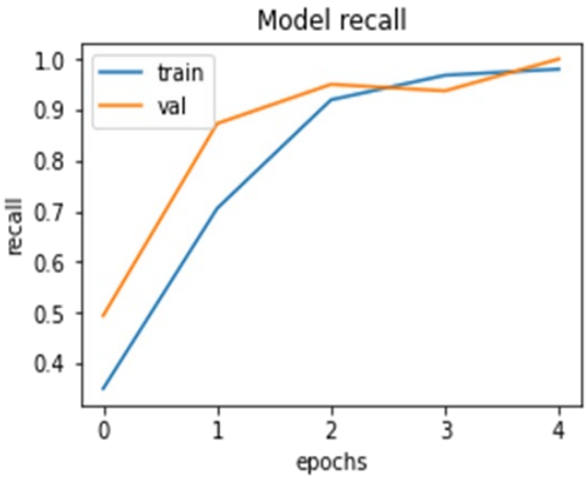
Precision = T P/ (T P + F P)

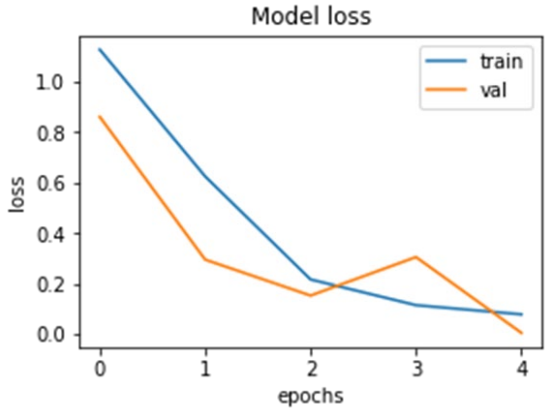


4.3 Recall

It is the ratio between actual positive cases that are predicted to be positive. In the context of this thesis, recall measures the fraction of leaf that are predicted as plant species name. Recall is calculated using the following formula.

Recall =T P/ (T P + F N)





4.4 F1 Score

It is also known as balanced F-score or F-measure. F1 score is a measure of accuracy of a model combining precision and recall. In the context of this thesis, a good F1 score shows that there are less false positives and false negatives. This shows that the model is correctly identifying plant species.

A model/algorithm is considered perfect if F1 score is 1. It is calculated using the following formula.

F1 = 2 × (Precision × Recall /Precision + Recall)
Accuracy: 0.9984025559105432
Precision: 0.9990234375
Recall: 0.9964285714285714
F1_score: 0.9977122020583142

4.4.1 Training time

Training time is metric used in this thesis to measure the time taken to train the selected machine learning algorithms on the dataset.

4.4.2 Prediction Speed

Speed is a metric used in this thesis to measure the time taken for the algorithms to process and detect obstacle.

4.4.3 Loss Function

Loss function, to perform feature matching between the ground truth and the output of segmentation network, optimizing also the network weights on features extracted at multiple resolutions rather than focusing just on the pixel level.

V. CONCLUSION

Classifying the plants species with their leaf using the algorithms of computer vision show concern on categorizing the plant images into its distinct groups. The classification of plants using digital leaf images are challenging due to their similarities in inter-class and intra-class, the possibility of complex background and variations in many parameters such as illumination and color. Thus, developing tools and solutions to analyze and interpret the patterns in the leaf images with significant results are essential. This project proposes an automated plant identification system, for identifying the plants species through their leaf. This task is accomplished using deep convolutional neural network to achieve higher accuracy. Image pre-processing, feature extraction and recognition are three main identification steps which are taken under consideration. Proposed CNN classifier learns the features of plants such as classification of leaves by using hidden layers like convolutional layer, max pooling layer, dropout layers and fully connected layers. The model acquires a knowledge related to features of Swedish leaf dataset in which 30plant classes are available, that helps to predict the correct category of unknown plant with accuracy of 97% and minimum losses. Result is slightly better than the previous work that analyzes 93.75% of accuracy.

REFERENCES

[1]. S. Mahajan, A. Raina, X.-Z. Gao, and A. K. Pandit, ``Plant recognition using morphological feature extraction and transfer learning over SVM and AdaBoost," Symmetry, vol. 13, no. 2, p. 356, Feb. 2021.
[2]. Beikmohammadi, K. Faez, and A. Motallebi, ``SWP-leaf NET: A novel multistage approach for plant leaf identification based on deep learning," 2020.

[3]. Muneer and S. M. Fati, ``Efficient and automated herbs classification approach based on shape and texture features using deep learning," IEEE Access, vol. 8, pp. 196747-196764, 2020.
[4]. J. W. Lee and Y. C. Yoon, ``Fine-grained plant identification using wide and deep learning model 1," in Proc. Int. Conf. Platform Technol. Service (PlatCon), Jan. 2019, pp. 1-5.
[5]. P. Jasitha, M. R. Dileep, and M. Divya, ``Venation based plant leaves classification using GoogLeNet and VGG," in Proc. 4th Int. Conf. Recent Trends Electron., Inf., Commun. Technol. (RTEICT), May 2019, pp. 715-719.
[6]. A. Kaya, A. S. Keceli, C. Catal, H. Y. Yalic, H. Temucin, andB. Tekinerdogan, ``Analysis of transfer learning for deep neural network-based plant classification models," Comput. Electron. Agricult., vol. 158, pp. 20-29, Mar. 2019.
[7]. V. Bodhwani, D. P. Acharjya, and U. Bodhwani, ``Deep residual networks for plant identification," Proc. Comput. Sci., vol. 152, pp. 186-194, Jan. 2019.
[8]. J. wei Tan, S.-W. Chang, S. Abdul-Kareem, H. J. Yap, and K.-T. Yong, “Deep learning for plant species classification using leaf vein morphometric,"IEEE/ACM Trans. Comput. Biol. Bioinf., vol. 17, no. 1, pp. 82-90, Jun. 2018.
[9]. Z. Ibrahim, N. Sabri, and N. N. A. Mangshor, ``Leaf recognition using texture features for herbal plant identification," Indonesian J. Electr. Eng.Comput. Sci., vol. 9, no. 1, pp. 152-156, 2018.
[10]. F. Chollet, ``Xception: Deep learning with depth wise separable convolutions,"in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 1251-1258.